Review of probability

To define a Probabilistic Model, we need

- A sample space $\Omega$: set of all possible outcomes of an experiment

  A subspace of the sample space is called event (the empty event is $\emptyset$)

- A probability law assigns to each event $A$ a number $P(A)$ such that
  - $0 \leq P(A) \leq 1$
  - The Probability of the entire sample space is $P(\Omega) = 1$
  - If $A$ and $B$ two disjoint events $A \cap B = \emptyset$ then
    \[ P(A \cup B) = P(A) + P(B) \]

From these axioms, we can deduce some important properties

- The Probability of the empty event is $P(\emptyset) = 0$
- If $A$ and $B$ are two events, then
  \[ P(A \cup B) = P(A) + P(B) - P(A \cap B) \]
Suppose we can completely partition $\Omega$ into $n$ disjoint events $A_1, A_2, \ldots, A_n$.

Any event $E$ can be decomposed by

$$P(E) = P(E \cap A_1) + P(E \cap A_2) + \ldots + P(E \cap A_n)$$

This can easily be understood from a Venn diagram.
Conditional probability

Appears when partial information about outcome is given.

Example: We roll a fair die, with the sample space \( \{1, 2, 3, 4, 5, 6\} \).

Q : What is the probability that the outcome is 4 ?

Since there are 6 different outcomes (all equally likely) and 4 is one of them, the answer is

\[
P(4) = \frac{\text{number of elements in } \{4\}}{\text{number of elements in } \Omega} = \frac{1}{6}
\]
Conditional probability

Appears when partial information about outcome is given.

Example: We roll a fair die, with the sample space \{1, 2, 3, 4, 5, 6\}.

Q : What is the probability that the outcome is 4 ?

Since there are 6 different outcomes (all equally likely) and 4 is one of them, the answer is

\[
P(4) = \frac{\text{number of elements in } \{4\}}{\text{number of elements in } \Omega} = \frac{1}{6}
\]

What happens if we know that the outcome is even ?

Q: What is the probability that the outcome is 4 given that the outcome is even?

\[
P(4|\text{even}) = P(4) = \frac{\text{number of elements in } \{4\} \cap \{\text{even}\}}{\text{number of elements in } \{\text{even}\}} = \frac{1}{3} = \frac{P(4 \cap \text{even})}{P(\text{even})} = \frac{1/6}{1/2}
\]
Conditional probability

Appears when partial information about outcome is given.

Example: We roll a fair die, with the sample space \{1, 2, 3, 4, 5, 6\}.

Q : What is the probability that the outcome is 4 ?

Since there are 6 different outcomes (all equally likely) and 4 is one of them, the answer is

\[
P(4) = \frac{\text{number of elements in } \{4\}}{\text{number of elements in } \Omega} = \frac{1}{6}
\]

What happens if we know that the outcome is even ?

Q: What is the probability that the outcome is 4 given that the outcome is even?

\[
P(4|\text{even}) = P(4) = \frac{\text{number of elements in } \{4\} \cap \{\text{even}\}}{\text{number of elements in } \{\text{even}\}} = \frac{1}{3} = \frac{P(4 \cap \text{even})}{P(\text{even})} = \frac{1/6}{1/2}
\]

For two events \(A\) and \(B\), we define the conditional probability

\[
P(A|B) = \frac{P(A \cap B)}{P(B)}
\]
⇒ Express $P(A|B)$ in terms of $P(B|A)$
⇒ Express $P(A|B)$ in terms of $P(B|A)$

The derivation is very simple, we just use $A \cap B = B \cap A \Rightarrow P(A \cap B) = P(B \cap A)$
Express $P(A|B)$ in terms of $P(B|A)$

The derivation is very simple, we just use $A \cap B = B \cap A \implies P(A \cap B) = P(B \cap A)$

and the definition of conditional probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \implies P(A \cap B) = P(A|B) \times P(B)$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \implies P(B \cap A) = P(B|A) \times P(A)$$
Bayes’ Theorem (Thomas Bayes 1701-1761)

⇒ Express $P(A|B)$ in terms of $P(B|A)$

The derivation is very simple, we just use $A \cap B = B \cap A \quad \Rightarrow P(A \cap B) = P(B \cap A)$

and the definition of conditional probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A|B) \times P(B)$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \Rightarrow P(B \cap A) = P(B|A) \times P(A)$$

Bayes’ theorem

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$
A radar is designed to detect aircraft. If an aircraft is present, it is detected with probability 0.99. When no aircraft is present, the radar generates an alarm probability 0.02 (false alarm). We assume that an aircraft is present with probability 0.05. If the radar generates an alarm, what is the probability than an aircraft is present?
A radar is designed to detect aircraft. If an aircraft is present, it is detected with probability 0.99. When no aircraft is present, the radar generates an alarm probability 0.02 (false alarm). We assume that an aircraft is present with probability 0.05. If the radar generates an alarm, what is the probability than an aircraft is present?

\[ P(\text{aircraft}|\text{alarm}) = \frac{P(\text{alarm}|\text{aircraft}) \times P(\text{aircraft})}{P(\text{alarm})} \]
A radar is designed to detect aircraft. If an aircraft is present, it is detected with probability 0.99. When no aircraft is present, the radar generates an alarm probability 0.02 (false alarm). We assume that an aircraft is present with probability 0.05. If the radar generates an alarm, what is the probability than an aircraft is present?

\[
P(\text{aircraft} | \text{alarm}) = \frac{P(\text{alarm} | \text{aircraft}) \times P(\text{aircraft})}{P(\text{alarm})}
\]

Since the set \{aircraft, no aircraft\} is a partition of \(\Omega\) we have

\[
P(\text{alarm}) = P(\text{alarm} \cap \text{aircraft}) + P(\text{alarm} \cap \text{no aircraft})
\]

\[
= P(\text{alarm} | \text{aircraft}) \times P(\text{aircraft}) + P(\text{alarm} | \text{no aircraft}) \times P(\text{no aircraft})
\]

Answer: 72%
First example: Detecting aircraft

A radar is designed to detect aircraft. If an aircraft is present, it is detected with probability 0.99. When no aircraft is present, the radar generates an alarm probability 0.02 (false alarm). We assume that an aircraft is present with probability 0.05. If the radar generates an alarm, what is the probability that an aircraft is present?

\[ P(\text{aircraft} | \text{alarm}) = \frac{P(\text{alarm} | \text{aircraft}) \times P(\text{aircraft})}{P(\text{alarm})} \]

Since the set \{aircraft, no aircraft\} is a partition of \( \Omega \) we have

\[ P(\text{alarm}) = P(\text{alarm} \cap \text{aircraft}) + P(\text{alarm} \cap \text{no aircraft}) \]
\[ = P(\text{alarm} | \text{aircraft}) \times P(\text{aircraft}) + P(\text{alarm} | \text{no aircraft}) \times P(\text{no aircraft}) \]

So we obtain

\[ P(\text{aircraft} | \text{alarm}) = \frac{P(\text{alarm} | \text{aircraft}) \times P(\text{aircraft})}{P(\text{alarm} | \text{aircraft}) \times P(\text{aircraft}) + P(\text{alarm} | \text{no aircraft}) \times P(\text{no aircraft})} \]
\[ = \frac{0.99 \times 0.05}{0.99 \times 0.05 + 0.02 \times 0.95} \sim 0.72 \]

Answer: 72%
Importance of Bayes’ Theorem

- There are many reasons why this theorem is important (for example in the interpretation of probability, in games theory, etc.)
There are many reasons why this theorem is important (for example in the interpretation of probability, in games theory, etc.)

An effect $E$ (alarm) can have different causes $C_1, C_2, C_3, \ldots$ (aircraft, false alarm).
There are many reasons why this theorem is important (for example in the interpretation of probability, in games theory, etc.)

An effect $E$ (alarm) can have different causes $C_1, C_2, C_3, \ldots$ (aircraft, false alarm).

- If the effect is observed, what is the probability that it is due to a given cause, say $C_1$?
There are many reasons why this theorem is important (for example in the interpretation of probability, in games theory, etc.)

An effect $E$ (alarm) can have different causes $C_1, C_2, C_3, \ldots$ (aircraft, false alarm).

- If the effect is observed, what is the probability that it is due to a given cause, say $C_1$?

  - $P(C_1|E) = \frac{P(E|C_1)}{P(E)} \times P(C_1)$

  - $P(C_1)$ is the initial degree of the belief in $C_1$ (prior)
  
  - $P(C_1|E)$ is the degree of the belief having accounted for $E$ (posterior)
There are many reasons why this theorem is important (for example in the interpretation of probability, in games theory, etc.)

An effect $E$ (alarm) can have different causes $C_1, C_2, C_3, \ldots$ (aircraft, false alarm).

- If the effect is observed, what is the probability that it is due to a given cause, say $C_1$?
  - $P(C_1|E) = \frac{P(E|C_1)}{P(E)} \times P(C_1)$
  - $P(C_1)$ is the initial degree of the belief in $C_1$ (prior)
  - $P(C_1|E)$ is the degree of the belief having accounted for $E$ (posterior)

In science, very often we have access to $P(A|B)$ (for example by some experiments) but what we really want to know is $P(B|A)$

We can then use Bayes’ theorem, provided we also know $P(A)$ and $P(B)$

$$P(B|A) = \frac{P(A|B)}{P(B)} \times P(A)$$

This is illustrated by the next example
Second example: The False Positive Puzzle

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease. The test results are assumed to be correct 95% of the time: if a person has the disease, the test results are positive with probability 0.95, and if the person does not have the disease, the test results are negative with probability 0.95.

Is this a reliable test? given that a person just tested positive, what is the probability of having the disease?
When developing the test in a lab, we take certain persons who are known to have the disease and run the test.

⇒ Therefore the scientists are measuring $\mathcal{P}(\text{positive}|\text{disease})$, $\mathcal{P}(\text{negative}|\text{disease})$.

We also perform the test on certain persons who do not have the disease, we then measure $\mathcal{P}(\text{positive}|\text{no disease})$, $\mathcal{P}(\text{negative}|\text{no disease})$. 
Second example: The False Positive Puzzle

- When developing the test in a lab, we take certain persons who are known to have the disease and run the test.

  ⇒ Therefore the scientists are measuring $P(\text{positive} | \text{disease})$, $P(\text{negative} | \text{disease})$.

  We also perform the test on certain persons who do not have the disease, we then measure $P(\text{positive} | \text{no disease})$, $P(\text{negative} | \text{no disease})$.

- Now we use the test in practise. If the results are positive, we want to know the probability that the patient really has the disease, so we want to know $P(\text{disease} | \text{positive})$.

  If the results are negative, what is the probability that the patient is infected (and that the test failed) ? $P(\text{disease} | \text{negative})$
Second example: The False Positive Puzzle

- When developing the test in a lab, we take certain persons who are known to have the disease and run the test.

  ⇒ Therefore the scientists are measuring $P(\text{positive}|\text{disease})$, $P(\text{negative}|\text{disease})$.

We also perform the test on certain persons who do not have the disease, we then measure $P(\text{positive}|\text{no disease})$, $P(\text{negative}|\text{no disease})$.

- Now we use the test in an practice. If the results are positive, we want to know the probability that the patient really has the disease, so we want to know $P(\text{disease}|\text{positive})$.

  If the results are negative, what is the probability that the patient is infected (and that the test failed) ? $P(\text{disease}|\text{negative})$

We want to find $P(\text{disease}|\text{positive})$ knowing $P(\text{positive}|\text{disease})$ ⇒ Bayes’ theorem
Second example: The False Positive Puzzle

\[ P(\text{disease}) = 0.001 \]

\[ P(\text{no disease}) = 0.999 \]

\[ P(\text{positive}|\text{disease}) = 0.95 \]

\[ P(\text{negative}|\text{disease}) = 0.05 \]

\[ P(\text{positive}|\text{no disease}) = 0.05 \]

\[ P(\text{negative}|\text{no disease}) = 0.95 \]

\[ P(\text{positive}|\text{positive}) = 0.95 \cdot 0.001 + 0.05 \cdot 0.999 \sim 0.01866 \]

The probability that the patient has the disease given that the test is positive is less than 2%!!
Second example: The False Positive Puzzle

Bayes’ theorem

\[
P(\text{disease}|\text{positive}) = \frac{P(\text{positive}|\text{disease}) \times P(\text{disease})}{P(\text{positive})}
\]

We need \(P(\text{positive})\), we use the fact that \text{disease} and \text{no disease} form a partition of \(\Omega\)

\[
P(\text{positive}) = P(\text{positive} \cap \text{disease}) + P(\text{positive} \cap \text{no disease})
\]

\[
= P(\text{positive}|\text{disease}) \times P(\text{disease}) + P(\text{positive}|\text{no disease}) \times P(\text{no disease})
\]
Second example: The False Positive Puzzle

Bayes’ theorem

\[ P(\text{disease}|\text{positive}) = \frac{P(\text{positive}|\text{disease}) \times P(\text{disease})}{P(\text{positive})} \]

We need \( P(\text{positive}) \), we use the fact that \textit{disease} and \textit{no disease} form a partition of \( \Omega \)

\[
P(\text{positive}) = P(\text{positive} \cap \text{disease}) + P(\text{positive} \cap \text{no disease}) = P(\text{positive}|\text{disease}) \times P(\text{disease}) + P(\text{positive}|\text{no disease}) \times P(\text{no disease})
\]

Therefore, using

\[ P(\text{positive}|\text{disease}) = 0.95, \ P(\text{positive}|\text{no disease}) = 0.05, \ P(\text{disease}) = 0.001, \] we find

\[
P(\text{disease}|\text{positive}) = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.05 \times 0.999} \sim 0.01866
\]
Second example: The False Positive Puzzle

Bayes’ theorem

\[ P(\text{disease}|\text{positive}) = \frac{P(\text{positive}|\text{disease}) \times P(\text{disease})}{P(\text{positive})} \]

We need \( P(\text{positive}) \), we use the fact that disease and no disease form a partition of \( \Omega \)

\[
P(\text{positive}) = P(\text{positive} \cap \text{disease}) + P(\text{positive} \cap \text{no disease})
\]

\[
= P(\text{positive}|\text{disease}) \times P(\text{disease}) + P(\text{positive}|\text{no disease}) \times P(\text{no disease})
\]

Therefore, using

\[ P(\text{positive}|\text{disease}) = 0.95, \quad P(\text{positive}|\text{no disease}) = 0.05, \quad P(\text{disease}) = 0.001, \quad \text{we find} \]

\[ P(\text{disease}|\text{positive}) = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.05 \times 0.999} \sim 0.01866 \]

The probability that the patient has the disease given that the test is positive is less than 2% !!
Solving the False Positive Puzzle

At first sight this result seems counter-intuitive, but there is a simple explanation.
Solving the False Positive Puzzle

At first sight this result seems counter-intuitive, but there is a simple explanation

The disease is rare: the probability of \textit{no disease} is a very close to one (0.999).

Therefore the probability of being positive and having the disease

\[ P(positive \cap disease) = 0.95 \times 0.001 = 0.00095 \]

is small compared to the probability of being a “false positive”

\[ P(positive \cap no disease) = 0.05 \times 0.999 = 0.04995 \]
Solving the False Positive Puzzle

At first sight this result seems counter-intuitive, but there is a simple explanation.

The disease is rare: the probability of no disease is a very close to one (0.999).

Therefore the probability of being positive and having the disease

\[ P(\text{positive} \cap \text{disease}) = 0.95 \times 0.001 = 0.00095 \]

is small compared to the probability of being a “false positive”

\[ P(\text{positive} \cap \text{no disease}) = 0.05 \times 0.999 = 0.04995 \]

The probability of being positive

\[ P(\text{positive}) = P(\text{positive} \cap \text{disease}) + P(\text{positive} \cap \text{no disease}) = 0.00095 + 0.04995 = 0.0509 \]

is largely dominated by \( P(\text{positive} \cap \text{no disease}) \)
Solving the False Positive Puzzle

At first sight this result seems counter-intuitive, but there is a simple explanation

The disease is rare: the probability of no disease is a very close to one (0.999).

Therefore the probability of being positive and having the disease

\[ P(\text{positive} \cap \text{disease}) = 0.95 \times 0.001 = 0.00095 \]

is small compared to the probability of being a “false positive”

\[ P(\text{positive} \cap \text{no disease}) = 0.05 \times 0.999 = 0.04995 \]

The probability of being positive

\[ P(\text{positive}) = P(\text{positive} \cap \text{disease}) + P(\text{positive} \cap \text{no disease}) \]

\[ = 0.00095 + 0.04995 = 0.0509 \]

is largely dominated by \( P(\text{positive} \cap \text{no disease}) \)

In other words: if somebody is tested positive, it is very likely that he is false positive

\[ P(\text{no disease}|\text{positive}) = \frac{0.04995}{0.0509} \sim 0.9811 \quad P(\text{disease}|\text{positive}) = \frac{0.00095}{0.0509} \sim 0.0186 \]
If we pick up a random person detected *positive*, most likely it is a *false positive*
P(negative) = 0.9491
P(positive) = 0.0509
P(disease|positive) = 0.0186
P(disease|negative) = 0.00005
P(no disease|positive) = 0.9811
P(no disease|negative) = 0.99995
Let us check by changing the numbers

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease if a person has the disease, the test results are positive with probability 0.95, and if the person does not have the disease, the test results are negative with probability 0.95.

given that a person just tested positive, what is the probability of having the disease ?

\[
P(\text{disease}|\text{positive}) = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.05 \times 0.999} \sim 0.01866
\]
Let us check by changing the numbers

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease.

If a person has the disease, the test results are positive with probability 0.95, and if the person does not have the disease, the test results are negative with probability 0.95.

given that a person just tested positive, what is the probability of having the disease?

\[
P(\text{disease} | \text{positive}) = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.05 \times 0.999} \sim 0.01866
\]
Let us check by changing the numbers

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease.

If a person has the disease, the test results are positive with probability 0.99, and if the person does not have the disease, the test results are negative with probability 0.95.

given that a person just tested positive, what is the probability of having the disease?

\[
P(\text{disease} | \text{positive}) = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.05 \times 0.999} \sim 0.01943
\]
Let us check by changing the numbers

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease

if a person has the disease, the test results are positive with probability 0.999,
and if the person does not have the disease, the test results are negative with probability 0.95.

given that a person just tested positive, what is the probability of having the disease ?

\[ P(\text{disease}|\text{positive}) = \frac{0.999 \times 0.001}{0.999 \times 0.001 + 0.05 \times 0.999} \sim 0.01961 \]
We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease.

If a person has the disease, the test results are positive with probability 1, and if the person does not have the disease, the test results are negative with probability 0.95.

Given that a person just tested positive, what is the probability of having the disease?

\[
P(\text{disease} | \text{positive}) = \frac{1 \times 0.001}{1 \times 0.001 + 0.05 \times 0.999} \sim 0.01963
\]
Let us check by changing the numbers

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease.

If a person has the disease, the test results are positive with probability 0.95, and if the person does not have the disease, the test results are negative with probability 0.95.

given that a person just tested positive, what is the probability of having the disease?

\[ P(\text{disease}|\text{positive}) = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.05 \times 0.999} \sim 0.01866 \]
Let us check by changing the numbers

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease.

If a person has the disease, the test results are positive with probability 0.95, and if the person does not have the disease, the test results are negative with probability 0.99.

Given that a person just tested positive, what is the probability of having the disease?

\[
P(\text{disease}|\text{positive}) = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.01 \times 0.999} \approx 0.08684
\]
Let us check by changing the numbers

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease.

if a person has the disease, the test results are positive with probability 0.95, and if the person does not have the disease, the test results are negative with probability 0.999.

given that a person just tested positive, what is the probability of having the disease?

\[
P(\text{disease}|\text{positive}) = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.001 \times 0.999} \approx 0.48743
\]
Let us check by changing the numbers

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease. If a person has the disease, the test results are positive with probability 0.95, and if the person does not have the disease, the test results are negative with probability 0.9999. Given that a person just tested positive, what is the probability of having the disease?

\[
P(\text{disease}|\text{positive}) = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.0001 \times 0.999} \sim 0.90834
\]
Let us check by changing the numbers

We want to know if a clinical test for a given rare disease is reliable. One person per 1,000 is affected by this disease
if a person has the disease, the test results are positive with probability 0.95,
and if the person does not have the disease, the test results are negative with probability 0.9999.
given that a person just tested positive, what is the probability of having the disease?

\[
P(\text{disease} | \text{positive}) = \frac{0.95 \times 0.001}{0.95 \times 0.001 + 0.0001 \times 0.999} \sim 0.90834
\]

Homework: change \(P(\text{disease})\) to 0.005, 0.01, 0.05, 0.1...
The last example shows the importance of a definition in probability: if a lab says a test has a success rates of 95%, one should ask if the number corresponds to $P(positive|disease)$ or to $P(disease|positive)$.

- Bayes theorem plays a crucial part in probability and statistics.
- We saw two simple applications.
- Its demonstration is very simple but the results can be surprising.
- Has a lot of important applications, in particular it “inverts” a probability diagram.
- An example: for a desired $P(disease|positive)$ what $P(positive|disease)$ should we require?